AD_____

Award Number: DAMD17-03-1-0529

TITLE: Using RNA Interference to Reveal Genetic Vulnerabilities in Human Cancer Cells

PRINCIPAL INVESTIGATOR: Despina Siolas
Gregory J. Hannon, Ph.D.

CONTRACTING ORGANIZATION: Cold Spring Harbor Laboratory
Cold Spring Harbor, NY 11724

REPORT DATE: July 2004

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2004	Annual Summary (1 Jul 2003 - 30 Jun 2004)

4. TITLE AND SUBTITLE
Using RNA Interference to Reveal Genetic Vulnerabilities in Human Cancer Cells

5. FUNDING NUMBERS
DAMD17-03-1-0529

6. AUTHOR(S)
Despina Siolas
Gregory J. Hannon, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Cold Spring Harbor Laboratory
Cold Spring Harbor, NY 11724

E-Mail: siolas@cshl.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Designing highly potent silencing triggers is key to successful application of RNAi in mammals. Recent studies suggested that assembly of RNAi effector complexes (RISC) is coupled to Dicer cleavage. We therefore examined whether transfection of optimized Dicer substrates might give an improved RNAi response. Dicer cleavage of chemically synthesized short hairpin RNAs (shRNAs) with 29 bp. stems and 2 nucleotide 3' overhangs produced predictable homogenous small RNAs comprising the 22 bases at the 3' end of the stem. Consequently, direct comparisons of synthetic siRNAs and shRNAs that yield the same small RNA became possible. We found synthetic, 29mer shRNAs to be more potent inducers of RNAi than siRNAs. Maximal inhibition of target genes can be achieved at lower concentrations and silencing at 24 hours is often greater. Overall, our results suggest an improved method for triggering RNAi in mammalian cancer cells that uses higher potency RNAi triggers. This remains a critical issue for both cell culture studies and for potential therapeutic use in vivo for cancer.

14. SUBJECT TERMS
Cancer, RNAi, gene silencing

15. NUMBER OF PAGES
28

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....1

SF 298.....2

Table of Contents.....3

Introduction.....4

Body.....6

Key Research Accomplishments.....7

Reportable Outcomes.....7

Conclusions.....7

References.....none

Appendices.....8

Introduction:

Many eukaryotic organisms respond to double stranded RNA (dsRNA) by initiating a sequence specific silencing pathway known as RNA intereference or RNAi. The ability to exploit RNAi as an experimental tool for cancer has evolved in lock-step with an elucidation of the underlying biochemical mechanism of this regulatory pathway. Due in part to the triggering of non-sequence specific responses by dsRNAs of greater than 30-50 nucleotides, the experimental use of RNAi in mammalian systems awaited a detailed understanding of the RNAi mechanism. Studies in a multitude of organisms led to the development of a methodology for experimentally programming the RNAi machinery in mammalian cells by direct delivery of chemically synthesized siRNAs. A second approach for triggering RNAi in mammalian cells came about using DNA vectors encoding short hairpin RNAs (shRNAs), modeled roughly after endogenous microRNAs.

The first task of my proposal was to determine the structural requirements and optimal expression strategies for short hairpin RNA (shRNA) that can be used in cultured mammalian cells. The ultimate goal is to use RNAi as a genetic tool to find molecular vulnerabilities unique to breast cancer cells. These vulnerabilities are potential chemotherapeutic targets that can be exploited to kill cancer cells.

In order to optimize our RNAi delivery method in cultured mammalian cancer cells, we investigated the intracellular processing of shRNAs. Remarkably, for both miRNAs and siRNAs, the two strands of the processed dsRNA are treated unequally. Cloning efforts in a variety of organisms yielded overwhelmingly one strand for each miRNA. A potential explanation for this outcome came from biochemical studies of siRNAs in *Drosophila* suggesting that relative thermodynamic instability at the 5' end of a given strand of the Dicer product favors its loading into RISC. This is in accord with analysis of predicted Dicer cleavage products of endogenous miRNAs and with studies of the efficacy of large numbers of siRNAs, which indicate greater suppression if the antisense strand (relative to the target mRNA) has an unstable 5' end. Recent reports have suggested that this loading might occur in a complex and be coordinated with Dicer cleavage. A possibility suggested by these mechanistic insights is that Dicer substrates might be

more efficiently incorporated into RISC than siRNAs. We therefore sought to understand how Dicer processes shRNAs in order to permit comparison of the efficiency of silencing triggers that are predicted to produce equivalent RISC enzymes.

Body:

The results of my work, which are attached in the paper "Synthetic shRNA as highly potent RNAi triggers" suggest that chemically synthesized, 29mer shRNAs are often substantially more effective triggers of RNAi than are siRNAs. A mechanistic explanation for this finding may lie in the fact that 29mer shRNAs are substrates for Dicer processing both *in vitro* and *in vivo*. Results from several laboratories have strongly suggested a model for assembly of the RNAi effector complex in which a multi-protein assembly containing Dicer and accessory proteins interacts with an Argonaute protein and actively loads one strand of the siRNA or miRNA into RISC. Such a model implies that Dicer substrates, derived from nuclear processing of pri-miRNAs or cytoplasmic delivery of pre-miRNA mimetics, might be loaded into RISC more effectively than siRNAs. My data support such a prediction, since it is not the hairpin structure of the synthetic RNA that determines its increased efficacy but the fact that the shRNA is a Dicer substrate that correlates with enhanced potency. In *Drosophila*, Dicer is also required for siRNAs to enter RISC, and similar data has been obtained in mammalian cells. Thus, it is possible that even siRNAs enter RISC via a Dicer-mediated assembly pathway and that our data simply reflect an increased affinity of Dicer for longer duplex substrates. Alternatively, hairpin RNAs, such as miRNA precursors, might interact with specific cellular proteins that facilitate delivery of these substrates to Dicer, whereas siRNAs might not benefit from such chaperones.

The use of RNAi in mammalian cells has the potential to change mammalian cell systems into powerful genetic systems. Our results suggest an improved method for triggering RNAi in mammalian cells that uses higher potency RNAi triggers. Mapping the predominant 22 nt. sequence that appears in RISC from each of these shRNAs now permits the combination of this more effective triggering method with rules for effective siRNA design. Now, I am beginning on the next phase of my project, to carry out genome wide screens using shRNAs targeting DNA replication and repair genes in MCF7 breast cancer cells for apoptosis and growth arrest. These screens will identify new chemotherapeutic targets that will provide us with new insight into the causes of breast cancer.

Key Research Accomplishments:

- Mapping the predominant 22 nt. sequence that appears in RISC from shRNAs allowing us to develop more potent RNAi triggers for mammalian cancer cell culture.
- Determining that chemically synthesized, 29 bp shRNAs are often substantially more effective triggers of RNAi than siRNAs

Reportable Outcomes:

Manuscripts:

Siolas, D, Lerner C, Burchard J, Ge W, Linsley P, Paddison P, Hannon G, Cleary MA. Synthetic shRNAs as highly potent RNAi triggers. Nature Biotechnology. Accepted for publication Nov. 4 2004. (see attached).

Conclusions:

I have finished the first Aim of my proposal to optimize the structural requirements and expression strategies for shRNAs. I have already begun to test out parameters for screens that will identify chemotherapeutic targets in MCF7 cells.

Synthetic shRNAs as highly potent RNAi triggers

Despina Siolas [1,2], Cara Lerner [3], Julja Burchard [3], Wei Ge[3], Peter Linsley [3], Patrick Paddison[2], Gregory J. Hannon [2*] and Michele A. Cleary [3*]

[1] Program in Genetics
Stony Brook University
Stony Brook, NY 11794

[2] Cold Spring Harbor Laboratory
Watson School of Biological Sciences
1 Bungtown Road
Cold Spring Harbor, NY 11724

[3]Rosetta Inpharmatics LLC,
A Wholly Owned Subsidiary of Merck & Co., Inc.
401 Terry Ave North
Seattle, Washington 98109, USA

* to whom correspondence should be addressed
hannon@cshl.edu
Michele_cleary@merck.com

Designing highly potent silencing triggers is key to successful application of RNAi in mammals. Recent studies suggested that assembly of RNAi effector complexes (RISC) is coupled to Dicer cleavage. We therefore examined whether transfection of optimized Dicer substrates might give an improved RNAi response. Dicer cleavage of chemically synthesized short hairpin RNAs (shRNAs) with 29 bp. stems and 2 nt. 3' overhangs produced predictable homogeneous small RNAs comprising the 22 bases at the 3' end of the stem. Consequently, direct comparisons of synthetic siRNAs and shRNAs that yield the same small RNA became possible. We found synthetic, 29mer shRNAs to be more potent inducers of RNAi than siRNAs. Maximal inhibition of target genes can be achieved at lower concentrations and silencing at 24 hours is often greater. These studies provide an improved methodology for triggering experimental silencing via the RNAi pathway.

Many eukaryotic organisms respond to double-stranded RNA (dsRNA) by activating a sequence-specific silencing pathway, known as RNA interference or RNAi. Initiation of RNAi occurs upon processing of double-stranded RNAs into ~22nt fragments, known as siRNAs [1-3], by an RNAseIII family nuclease, Dicer [4]. These small RNAs are used as guides for selection and cleavage of complementary mRNAs through their incorporation into an effector complex, RISC,[2, 3, 5, 6] whose catalytic subunit, Argonaute 2 has recently been identified [7, 8]. These mechanistic insights led to the development of a methodology for experimentally programming the RNAi machinery in mammalian cells by directly transfecting chemically synthesized siRNA duplexes of ~21 nt. consisting of 19 paired bases with 2 nucleotide 3' overhangs to produce a transient silencing response [9-11].

In many organisms, the RNAi machinery also serves an effector function for endogenous, non-coding RNAs, known as microRNAs (miRNAs) (reviewed in [12]). MiRNAs are initially generated as long primary transcripts (pri-miRNA), which are cleaved in the nucleus by another RNAseIII family nuclease, Drosha [13]. The liberated pre-miRNAs are exported to the cytoplasm where Dicer performs a second cleavage to produce small RNAs that are loaded into RISC [14-17]. In the case of miRNAs, the cleavage sites are specific, and most often a single, discrete sequence is liberated from the precursor (reviewed in [12]). These discoveries prompted the development of a second approach for triggering RNAi in mammalian cells using DNA vectors encoding short hairpin RNAs (shRNAs), modeled roughly after endogenous microRNAs [18-23].

Remarkably, for both miRNAs and siRNAs, the two strands of the processed dsRNA are treated unequally. Cloning efforts in a variety of organisms yielded overwhelmingly one strand for each miRNA [24-26]. A potential explanation for this outcome came from biochemical studies of siRNAs in *Drosophila* suggesting that relative thermodynamic instability at the 5' end of a given strand of the Dicer product favors its loading into RISC [27]. This is in accord with analysis of predicted Dicer cleavage products of endogenous miRNAs [28, 29] and with studies of the efficacy of large numbers of siRNAs, which indicate greater suppression if the antisense strand (relative to the target mRNA) has an unstable 5' end [29]. Recent reports have suggested that this loading might occur in a complex and be coordinated with Dicer cleavage [30-32]. A possibility suggested by these mechanistic insights is that Dicer substrates might be more efficiently incorporated into RISC than siRNAs. We therefore sought to understand how Dicer processes shRNAs in order to permit comparison of the efficiency of silencing triggers that are predicted to produce equivalent RISC enzymes.

Dicer cleaves a predominant small RNA from the end of each shRNA

We began by producing ~70 chemically synthesized shRNAs, targeting various endogenous genes and reporters. We focused on a detailed analysis of

one set of four shRNAs that target firefly luciferase (Fig 1A). The individual species differed in two distinct ways. First, the stems of the shRNAs were either 19 or 29 bp in length. These sizes were chosen to reflect the two stem sizes most commonly used for vector-expressed shRNAs. Second, each shRNA either contained or lacked a 2 nt. 3' overhang, identical to that produced by processing of pri-miRNAs by Drosha. Each species was end-labeled by enzymatic phosphorylation and incubated with recombinant human Dicer. The 29mer shRNA bearing the 3' overhang was converted almost quantitatively into a 22 nt. product by Dicer (Fig. 1B). In contrast, the 29mer shRNA that lacked the overhang generated very little discrete 22 nt. labeled product, despite a Dicer-dependent depletion of the starting material. Neither 19mer shRNA was cleaved to a significant extent by the Dicer enzyme. This result was not due to the lack of double stranded structure in the 19mer shRNAs, as all shRNA substrates were efficiently cleaved by bacterial RNAseIII (Supplemental Figure 1). These results suggested that the shRNAs bearing a 3'overhang produced predominantly one specific and unique small RNA product, while a blunt ended hairpin was processed into a range of products. This hypothesis was consistent with parallel analysis of identical shRNA substrates that were produced by *in vitro* transcription with T7 polymerase and uniformly labeled (Fig. 1C). Cleavage of the uniformly labeled hairpin with an overhang resulted in products accumulating from 29mer shRNAs both with and without overhangs. In the latter case, the products did not accumulate to the same degree as seen with the overhang-containing material. Additionally, shRNAs with overhangs yielded products of two discrete sizes (21 and 22 nt.). Considered together, our results suggest that Dicer requires a minimum stem length for efficient cleavage. Furthermore, they are consistent with a hypothesis that the presence of a correct 3' overhang enhances the efficiency and specificity of cleavage, directing Dicer to cut ~22 nt. from the end of the substrate.

A number of previous studies have suggested that Dicer might function as an end-recognizing endonuclease, without positing a role for the 3' overhang. Blocking of the ends of dsRNAs using either fold-back structures or chimeric RNA-DNA hybrids attenuated, but did not abolish, the ability of human Dicer to generate siRNAs [33]. Lund and colleagues suggested that Dicer cleaved ~22 nt from the blunt end of an extended pre-miRNA, designed in part to mimic a pri-miRNA (see [34]). Structural analysis of the Argonaute 2 PAZ domain suggested that it engages very short (~2-3 nt.) stretches of the 3' ends of single-stranded RNAs [35-39]. This led Song and colleagues to propose a model in which the 3' overhangs of pre-miRNAs, generated by Drosha cleavage, would serve as an important recognition and specificity determinant for subsequent processing by Dicer [38]. The results presented here are consistent with this model and suggest further that the 3' overhang aids in determining the specificity of cleavage, directing processing to a site 22 nt. from the 3' end of the substrate. These findings are in full accord with a recently published model for Dicer action [40].

To validate our biochemical analysis, we also mapped the position of Dicer cleavage *in vivo* using primer extension. Precursors were transfected into cells, and the processed form of each was isolated by virtue of its co-immunoprecipitation with co-expressed myc-tagged human Argonaute proteins, Ago1 and Ago2. The 29mer shRNA with an overhang gave rise to a relatively discrete product of 20 nt. as predicted for a cleavage 22 nt. from the 3'end of the substrate. Primer extension suggested identical cleavage specificities upon exposure of shRNAs to Dicer *in vitro* and in living cells (Fig 2A). Control experiments using a luciferase 29mer shRNA alone (without myc-tagged Ago1 or Ago2 expression) or cells transfected with myc-tagged Ago1 or Ago2 alone (no shRNA) did not yield extension products (Fig 2B).

Although the inability of Dicer to effectively cleave shRNAs with 19 bp. stems may seem at odds with the effective use of such structures for triggering RNAi using vector-based expression, there is presently no evidence that these RNAs require Dicer for their action. Indeed, our results using RNAi to deplete Dicer from cells suggests a strong dependence of silencing on Dicer for shRNAs with 29 bp. stems with a reduced or lack of dependency on Dicer for shRNAs with 19 bp. stems (not shown). However, 19mer shRNA do enter RISC. Human 293 cells that constitutively express Ago 1 were transfected with siRNAs, 29mer shRNAs or 19mer shRNAs. RISC was recovered by immunoprecipitation and associated RNAs were examined by Northern blotting. (Supplementary Figure 2) The 29mer shRNA with an overhang and the 22mer siRNA both entered RISC, giving 22 nt. small RNAs. The 19mer shRNA also entered RISC but gave 2 distinct small RNAs of 21 and 23 nt. While we do not understand the mechanistic basis for this observation, it may reflect Dicer-independent cleavage of the 19mer shRNA in the loop by a single-strand specific ribonuclease.

shRNAs are generally more effective than siRNAs

Since we could predict which single, specific 22 nt. sequence would be incorporated into RISC from a given shRNA, we could directly compare the silencing efficiency of shRNAs and siRNAs. Toward this goal, we selected 43 sequences targeting a total of 5 genes (3-9 sequences per gene). For each sequence, we synthesized a 21mer siRNA (19 bp stem) and shRNAs with 19 or 29 bp. stems that were predicted to give Dicer products that were either identical to their corresponding siRNAs or that differed by the addition of one 3' nucleotide homologous to the target. Importantly, each was predicted to give precisely the same 5' end following cleavage of a 22mer RNA from the shRNA (N in Supplementary Figure 3). Sequences for siRNAs are provided in Supplemental Table 1. Each RNA species was transfected into HeLa cells at a relatively high concentration (100 nM). The level of suppression was determined by semi-quantitative RT-PCR of RNA from HeLa cells 24 hours after transfection and the performance of each shRNA compared with the performance of the corresponding siRNA. Studies assessing siRNAs and 19mer shRNAs revealed that there was little difference in silencing at 24 hours with these species (Fig.

3A). A comparison of siRNAs with shRNAs having 29 bp. stems gave a different result. Clustering of the data points above the diagonal indicated consistently better inhibition with the 29mer shRNAs (Fig. 3B). As predicted by the aforementioned results, direct comparisons of shRNAs with 19 and 29 basepair stems indicated a greater overall effect with latter structure (Fig 3C).

The generally better inhibition with 29mer shRNAs at the high dose used for these studies led us to investigate the potency of these silencing triggers as compared with siRNAs. Seventeen complete sets comprising an siRNA, a 19mer shRNA and a 29mer shRNA were examined for suppression in titration experiments. In all cases, the 19mer shRNAs performed as well as or worse than the corresponding siRNAs. In contrast, 29mer shRNAs exceeded the performance of siRNAs in the majority of cases. In most cases, the 29mer shRNAs showed greater inhibition at the maximal dose; however even when this inhibition at the maximal dose did not differ much from the siRNA or 19mer shRNA the efficacy of the 29mer at lower concentrations was significantly better. The dose response curves for four representative sets of RNAs are shown in Figure 3D-H.

Consistent with our results for most of the RNA sets tested, in the case of MAPK14, KIF14 and KIF11, the maximal level of suppression for the 29mer shRNA was approximately two-fold greater than the maximal level of suppression for the corresponding siRNA (Fig. 3 E, G, H). More importantly, in some cases, the amount of RNA required to achieve maximal inhibition was up to 20-fold lower with 29mer shRNA than with a similar 21mer siRNA. The increase in potency for 29mer shRNA versus the other two RNA species may reflect the entry of these RNAs into the RNAi pathway as natural intermediates and may explain their greater efficacy when delivered from vectors [19].

siRNAs and shRNAs give similar profiles of off-target effects at saturation

Microarray analysis has revealed down-regulation of many non-targeted transcripts following transfection of siRNAs into HeLa cells [41]. Notably, these gene expression signatures differed between different siRNAs targeting the same gene. Many of the "off target" transcripts contained sites of partial identity to the individual siRNA, possibly explaining the source of the effects. To examine potential off-target effects of synthetic shRNAs, we compared shRNA signatures with those of siRNAs derived from the same target sequence. Using microarray gene expression profiling, we obtained a genome-wide view of transcript suppression. Figure 4 shows a two-dimensional clustering analysis of the signatures produced in HeLa cells 24 hours after transfection of 19mer and 29mer shRNAs compared with those generated by corresponding siRNAs. As indicated by the dendrogram on the vertical axis, each set of three RNAs derived from the same core sequence was accurately clustered. Furthermore, in all but 2 of 7 cases, although the19mer shRNAs produced signatures similar to those of the corresponding siRNAs, the signatures of the 29mer shRNAs and the siRNAs

were more closely related. Note, that in one of the 2 cases in which the 19mer shRNA and the siRNA clustered more closely (MAPK14-1), these two RNA species did not significantly silence the target gene, whereas the 29mer shRNA did. The agreement between the signatures of 29mer shRNAs and siRNAs is consistent with precise processing of the shRNA to generate a single siRNA rather than a random sampling of the hairpin stem by Dicer. The overall smaller signature sizes of the 19mer shRNA and the basis of their divergence from the signature of the corresponding siRNA is presently unclear. However, our goal was not to extensively analyze off target effects potentially associated with these shRNAs.

Discussion

Considered together, our results suggest that chemically synthesized, 29mer shRNAs are often substantially more effective triggers of RNAi than are siRNAs. A mechanistic explanation for this finding may lie in the fact that 29mer shRNAs are substrates for Dicer processing both *in vitro* and *in vivo*. We originally suggested that siRNAs might be passed from Dicer to RISC in a solid-state reaction on the basis of an interaction between Dicer and Argonaute2 in *Drosophila* S2 cell extracts [5]. More recently, results from several laboratories have strongly suggested a model for assembly of the RNAi effector complex in which a multi-protein assembly containing Dicer and accessory proteins interacts with an Argonaute protein and actively loads one strand of the siRNA or miRNA into RISC [30-32]. Such a model implies that Dicer substrates, derived from nuclear processing of pri-miRNAs or cytoplasmic delivery of pre-miRNA mimetics, might be loaded into RISC more effectively than siRNAs. Our data support such a prediction, since it is not the hairpin structure of the synthetic RNA that determines its increased efficacy but the fact that the shRNA is a Dicer substrate that correlates with enhanced potency (see also accompanying paper by Rossi and colleagues). In *Drosophila*, Dicer is also required for siRNAs to enter RISC, and similar data has been obtained in mammalian cells [30, 42]. Thus, it is possible that even siRNAs enter RISC via a Dicer-mediated assembly pathway and that our data simply reflect an increased affinity of Dicer for longer duplex substrates. Alternatively, hairpin RNAs, such as miRNA precursors, might interact with specific cellular proteins that facilitate delivery of these substrates to Dicer, whereas siRNAs might not benefit from such chaperones. Overall, our results suggest an improved method for triggering RNAi in mammalian cells that uses higher potency RNAi triggers. This remains a critical issue for both cell culture studies and for potential therapeutic use in vivo. Mapping the predominant 22 nt. sequence that appears in RISC from each of these shRNAs now permits the combination of this more effective triggering method with rules for effective siRNA design.

Methods

RNA sequence design

Each set of RNAs began with the choice of a single 19mer sequence. These 19mers were used directly to create siRNAs. To create shRNAs with 19mer stems, we appended a 4-base loop (either CCAA or UUGG) to the end of the 19mer sense strand target sequence followed by the 19mer complementary sequence and a UU overhang. We tested a variety of loop sequences and have noted no significant influence of the sequences examined on the performance of triggers. To create 29mer stems, we increased the length of the 19mer target sequence by adding 1 base upstream and 9 bases downstream from the target region and used the same loop sequence and UU overhang. All synthetic RNA molecules used in this study were purchased from Dharmacon.

Dicer processing

RNA hairpins corresponding to luciferase were end-labeled with [γ-32P] ATP and T4 Polynucleotide kinase. 0.1 pmoles of RNA were then processed with 2 units of Dicer (Stratagene) at 37 °C for 2 hours. Reaction products were trizol extracted, isopropranol precipitated, run on an 18% polyacrylamide, 8M urea denaturing gel. For RNaseIII digestion, 0.1 pmoles were digested with 1 unit of E. coli RNase III (NEB) for 30 minutes at 37 °C and analyzed as described above. Uniformly labeled hairpins were produced using a T7 Megashortscript kit (Ambion) with [α- 32P] UTP and then incubated with dicer as indicated above. For primer extension analysis, hairpins were processed with Dicer at 37 °C for 2 hours, followed by heat inactivation of the enzyme. The sequence of the primer is the first 16 nucleotide sequence from the 5' end of the hairpin: 5' AGTTGCGCCCGCGAAC 3'. DNA primers were 5' labeled with PNK and annealed to 0.05 pmole of RNA as follows : 95 °C for one minute, 10 minutes at 50 °C and then 1 min on ice. Extensions were carried out at 42 °C for 1 hour using MoMLV reverse transcriptase. Products were analyzed by electrophoresis on a 8M Urea/20% polyacrylamide gel. For analysis of *in vivo* processing, LinxA cells were transfected in 10 cm plates using Mirus TKO (10 ug hairpin RNA) or Mirus LT4 reagent for DNA transfection (12 ug of tagged Ago1/Ago 2 DNA; J. Liu, unpublished). 293 cells constitutively expressing Ago1 were utilized for Northern blot experiments. Cells were lysed and immunoprecipitated after 48 hours using with myc Antibody (9E14) Antibody. IPs were washed 3x in lysis buffer and treated with DNase for 15 minutes. Immunoprecipitates were then primer extended as described above.

siRNA and shRNA Transfections and mRNA Quantitation

HeLa cells were transfected in 96-well plates by use of Oligofectamine (Invitrogen) with the final nanomolar concentrations of each synthetic RNA indicated in the graphs. RNA quantitation was performed by Real-time PCR, using appropriate Applied Biosystems TaqMan™ primer probe sets 24 hours

after RNA transfection, and the percent mRNA remaining was compared with cells treated with transfection reagent alone.

Microarray Gene Expression Profiling

HeLa cells were transfected in 6-well plates with 100 nM final concentration of the appropriate RNA by use of Oligofectamine (according to the manufacturer's instructions). RNA from transfected cells was hybridized competitively with RNA from mock-transfected cells (treated with transfection reagent in the absence of synthetic RNA). Total RNA was purified by the Qiagen RNeasy kit, and processed as described previously [43]for hybridization to microarrays containing oligonucleotides corresponding to approximately 21,000 human genes. Ratio hybridizations were performed with fluorescent label reversal to eliminate dye bias. Microarrays were purchased from Agilent Technologies. Error models have been described previously [43]. Data were analyzed using Rosetta Resolver™ software.

Acknowledgements

G.J.H. is supported by an Innovator Award from the U.S. Army Breast Cancer Research Program. This work was also supported by a grant from the NIH (GJH). D.S. is supported by a predoctoral fellowship from the US Army Breast Cancer Research Program. We thank the Rosetta Gene Expression Laboratory for microarray RNA processing and hybridizations.

Figure Legends

Figure 1. *In vitro* processing of 29 bp. shRNAs by Dicer generates a predominant siRNA from the end of each short hairpin. **A.** The set of shRNAs containing 19 or 29 bp. stems and either bearing or lacking a 2 nt. 3'overhang is depicted schematically. For reference the 29 nt sequence from luciferase (top,blue) strand is given. The presumed cleavage sites (as predicted by analysis of Dicer processing products) are indicated in green and by the arrows. **B.** *In vitro* Dicer processing of shRNAs. 5' end-labeled substrates as depicted in **A** were incubated either in the presence or absence of recombinant human Dicer (as indicated). Processing of a 500 bp. blunt-ended dsRNA is shown for comparison. Markers are end-labeled, single-stranded, synthetic RNA oligonucleotides. **C.** Uniformly labeled shRNAs with structures as indicated in Fig. 1A were processed by Dicer to produce a small RNA product (as indicated). A long 500 bp. blunt-ended dsRNA is processed and shown for comparison.

Figure 2. Primer extension analysis reveals similar small RNAs are generated by Dicer processing *in vitro* or *in vivo*. A. Primer extension was used to analyze products from processing of overhung 29 mer shRNAs *in vivo*.

Total RNAs were extended with a specific primer that yields a 20 base product if cleavage occurs 22 bases from the 3' end of the overhang-containing RNA (see Fig 1A). For comparison, extensions of *in vitro* processed material are also shown. Lanes labeled siRNA are extensions of synthetic RNAs corresponding to predicted siRNAs that would be released by cleavage 21 or 22 nt. from the 3' end of the overhung precursor. Observation of extension products depends entirely on the inclusion of RT (indicated). The * indicates the specific extension product. Markers are phosphorylated, synthetic DNA oligonucleotides. **B.** Total RNA from control transfections, as indicated, did not show a primer extension product. The same primer was used for all extensions and is compatible with all RNAs. Controls with each RNA, as indicated, lacked a co-expressed tagged Ago protein, making it impossible to recover small RNAs in the immunoprecipitates. Controls labeled Ago1 or Ago2 lacked co-transfected target RNAs.

Figure 3. Gene suppression by shRNAs is comparable to or more effective than that achieved by siRNAs targeting the same sequences. A,B, C. mRNA suppression levels achieved by 43 siRNAs targeting 6 different genes were compared with levels achieved by 19mer or 29mer shRNAs derived from the same target sequences. 19mer and 29mer shRNAs were also directly compared. All RNAs were transfected at a final concentration of 100 nM. Values indicated on the *X* and *Y* axes reflect the percentage of mRNA remaining in HeLa cells 24 hours after RNA transfection compared with cells treated with transfection reagent alone. **D-H.** Four representative sets of siRNA and 19mer and 29mer shRNAs were used in dose-response analysis to compare the potency of representative RNAi triggers targeting 4 genes. Comparisons of relative suppression at the maximal dose are shown for reference in panel **D.** Titration curves were also performed reporting the percent of target mRNA remaining (*Y* axis) graphed from data derived from transfections at 1.56, 6.25, 25, and 100 nM final concentrations of RNA (*X* axis). Percent remaining RNA was determined by semi-quantitative RT-PCR. Gene targets were MAPK14, KIF 11, IGF1R and KIF14. (Sequences used were MAPK14-4, KIF11-6, IGF1R-1, KIF14-1 as in Suppl. Table 1.) (Blue diamonds: 21mer siRNAs; pink squares: 19mer shRNAs; green triangles: 29mer shRNAs). Red lines indicate the concentration of 29mer shRNA that gives the level of inhibition achieved by 100nM siRNA.

Figure 4. Microarray profiling reveals gene expression profiles are more similar between 29mer shRNAs and corresponding siRNAs than between siRNAs and 19mer shRNAs. 19mer and 29mer shRNAs and siRNAs designed for seven different target sequences within the coding region of MAPK14 were tested for gene silencing 24 hours after transfection into HeLa cells. Each row of the heat map reports the gene expression signature resulting from transfection of an individual RNA. Two-dimensional clustering of the data groups RNAs (vertical axis dendrogram) and regulated genes (horizontal axis dendrogram) according to signature similarities. Data shown represent genes that display at least a 2-fold change in expression level (p value<0.01 and $\log_{10}$ intensity >1) relative to mock-

transfected cells. Green indicates decreased expression relative to mock transfection whereas red indicates elevated expression. Black indicates no change and gray indicates data with a p value >0.01. The red arrow indicates MAPK14.

Supplementary Figure 1. The set of shRNAs containing 19 or 29 bp. stems coding a luciferase sequence and either bearing or lacking a 2 nt. 3'overhang were incubated with bacterial RNase III to verify their double-stranded nature.

Supplementary Figure 2. Northern blotting indicates that siRNAs and 19mer and 29mer shRNAs all give rise to RISC. Each RNA species was transfected into HeLa cells constitutively expressing tagged Ago1 protein. Small RNAs in RISC were detected by Northern blotting of Ago immunoprecipitates.

Supplementary Figure 3. Structures of synthetic RNAs used for comparing siRNA and shRNA. A total of 43 sequences were used with a matching set containing a 21mer siRNA (19 bp stem) and shRNAs with 19 or 29 bp. stems that were predicted to give Dicer products that were either identical to their corresponding siRNAs or that differed by the addition of one 3' nucleotide homologous to the target. Each was predicted to give precisely the same 5' end following cleavage of a 22mer RNA from the shRNA.

Supplementary Table 1. Sequences of the siRNAs used in this study

Gene	Accession number	Target sequence ID	Target sequence
IGF1R	NM_000875	IGF1R-1	GGAUGCACCAUCUUCAAGG
IGF1R	NM_000875	IGF1R-2	GACAAAAUCCCCAUCAGGA
IGF1R	NM_000875	IGF1R-3	ACCGCAAAGUCUUUGAGAA
IGF1R	NM_000875	IGF1R-4	GUCCUGACAUGCUGUUUGA
IGF1R	NM_000875	IGF1R-5	GACCACCAUCAACAAUGAG
IGF1R	NM_000875	IGF1R-6	CAAAUUAUGUGUUUCCGAA
IGF1R	NM_000875	IGF1R-7	CGCAUGUGCUGGCAGUAUA
IGF1R	NM_000875	IGF1R-8	CCGAAGAUUUCACAGUCAA
IGF1R	NM_000875	IGF1R-9	ACCAUUGAUUCUGUUACUU
KIF11	NM_004523	KIF11-1	CUGACAAGAGCUCAAGGAA
KIF11	NM_004523	KIF11-2	CGUUCUGGAGCUGUUGAUA
KIF11	NM_004523	KIF11-3	GAGCCCAGAUCAACCUUUA
KIF11	NM_004523	KIF11-4	GGCAUUAACACACUGGAGA
KIF11	NM_004523	KIF11-5	GAUGGCAGCUCAAAGCAAA
KIF11	NM_004523	KIF11-6	CAGCAGAAAUCUAAGGAUA
KIF11	NM_004523	KIF11-7	GACCUGUGCCUUUUAGAGA
KIF11	NM_004523	KIF11-8	AAAGGACAACUGCAGCUAC
KIF11	NM_004523	KIF11-9	GACUUCAUUGACAGUGGCC
KIF14	NM_014875	KIF14-1	CAGGGAUGCUGUUUGGAUA
KIF14	NM_014875	KIF14-2	ACUGACAACAAAGUGCAGC
KIF14	NM_014875	KIF14-3	AAACUGGGAGGCUACUUAC
KIF14	NM_014875	KIF14-4	CACUGAAUGUGGGAGGUGA
KIF14	NM_014875	KIF14-5	GUCUGGGUGGAAAUUCAAA
KIF14	NM_014875	KIF14-6	CAUCUUUGCUGAAUCGAAA
KIF14	NM_014875	KIF14-7	GGGAUUGACGGCAGUAAGA
KIF14	NM_014875	KIF14-8	CAGGUAAAGUCAGAGACAU

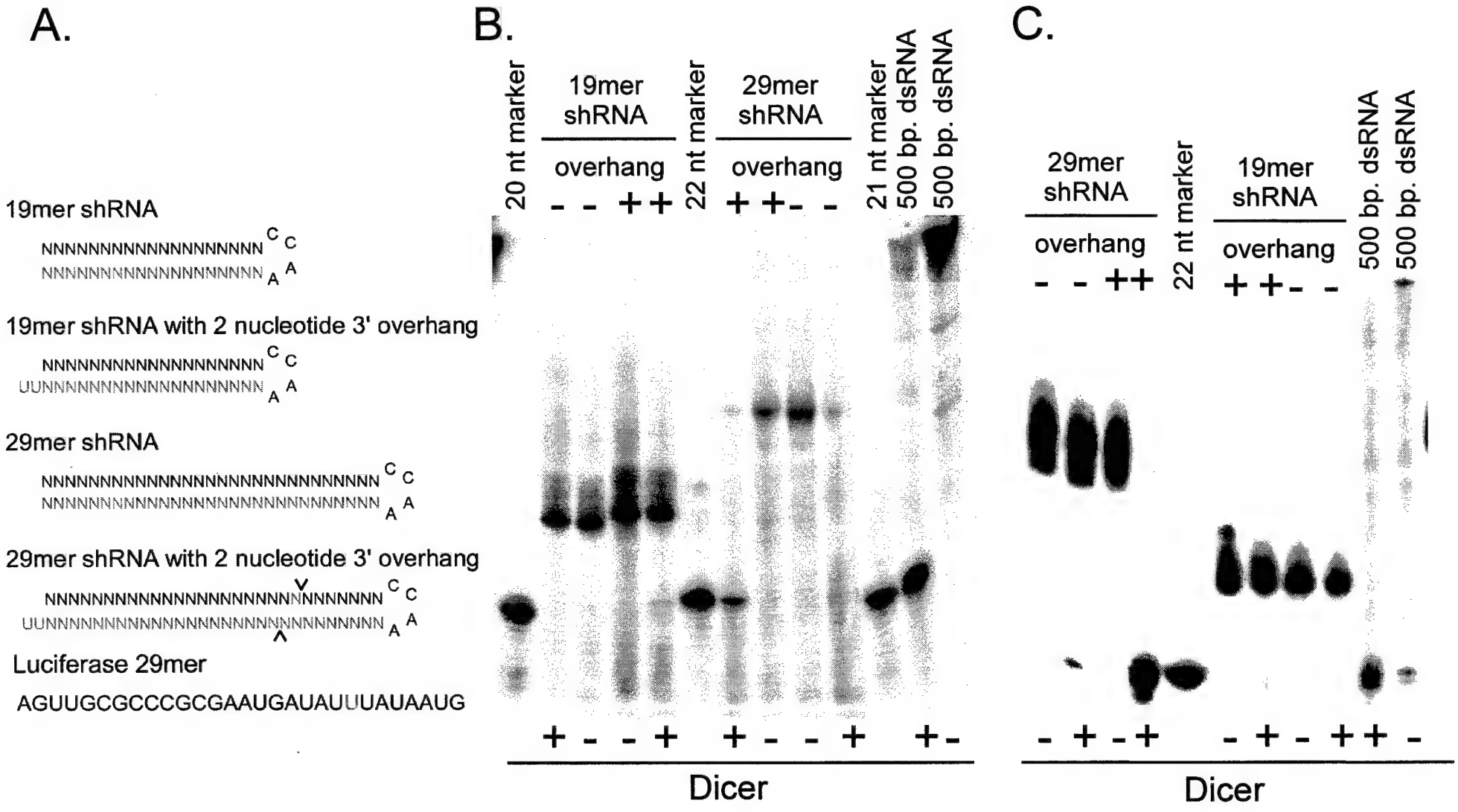
KIF14	NM_014875	KIF14-9	CUCACAUUGUCCACCAGGA
MAPK14	NM_139012	MAPK14-1	AAUAUCCUCAGGGGUGGAG
MAPK14	NM_139012	MAPK14-2	GUGCCUCUUGUUGCAGAGA
MAPK14	NM_139012	MAPK14-3	GAAGCUCUCCAGACCAUUU
MAPK14	NM_001315	MAPK14-4	CUCCUGAGAUCAUGCUGAA
MAPK14	NM_001315	MAPK14-5	GCUGUUGACUGGAAGAACA
MAPK14	NM_001315	MAPK14-6	GGAAUUCAAUGAUGUGUAU
MAPK14	NM_001315	MAPK14-7	CCAUUUCAGUCCAUCAUUC
PLK	NM_005030	PLK-1	CCCUGUGUGGGACUCCUAA
PLK	NM_005030	PLK-2	CCGAGUUAUUCAUCGAGAC
PLK	NM_005030	PLK-3	GUUCUUUACUUCUGGCUAU
PLK	NM_005030	PLK-4	CGCCUCAUCCUCUACAAUG
PLK	NM_005030	PLK-5	AAGAGACCUACCUCCGGAU
PLK	NM_005030	PLK-6	GGUGUUCGCGGGCAAGAUU
PLK	NM_005030	PLK-7	CUCCUUAAAUAUUUCCGCA
PLK	NM_005030	PLK-8	AAGAAGAACCAGUGGUUCG
PLK	NM_005030	PLK-9	CUGAGCCUGAGGCCCGAUA

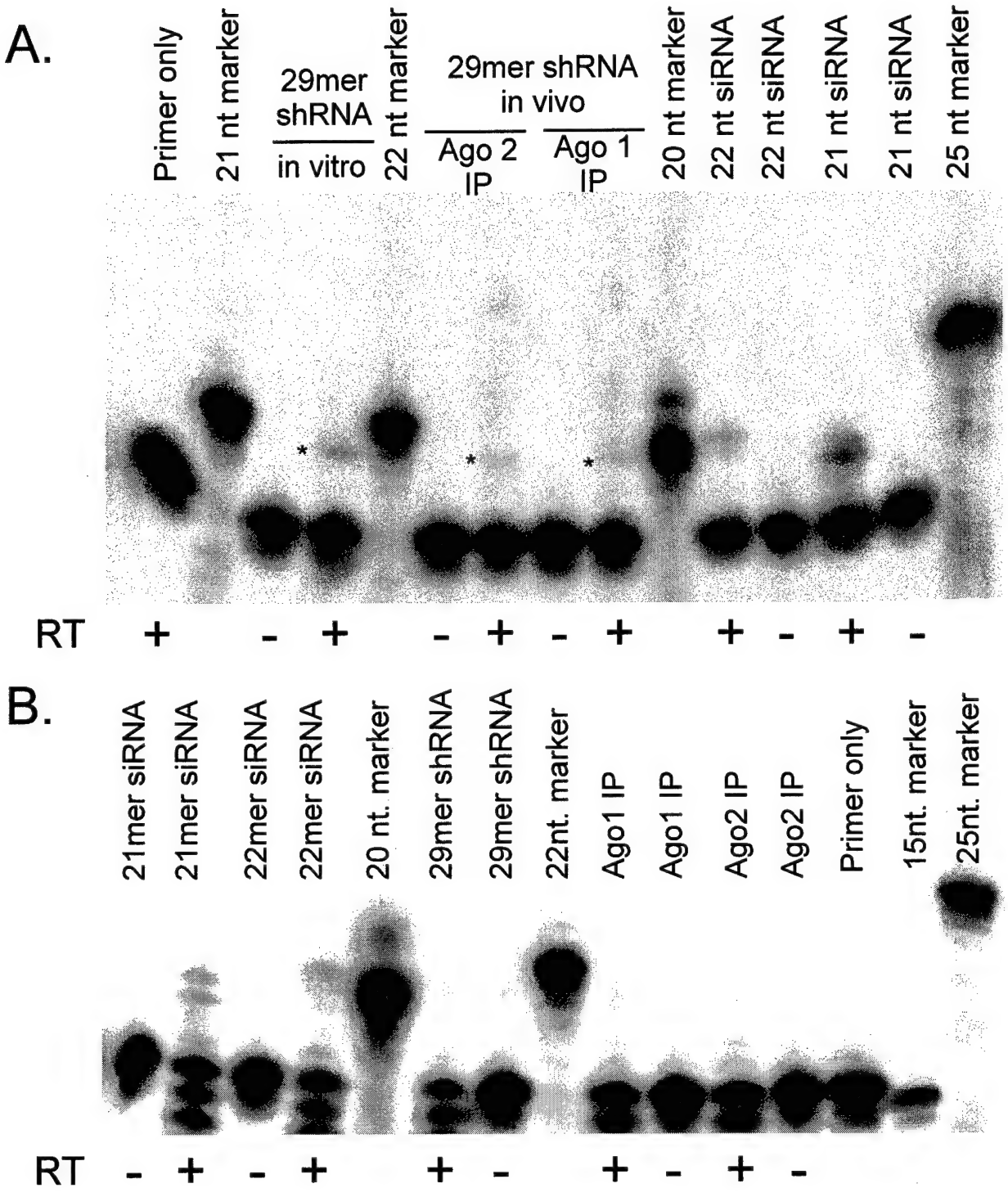
Literature Cited

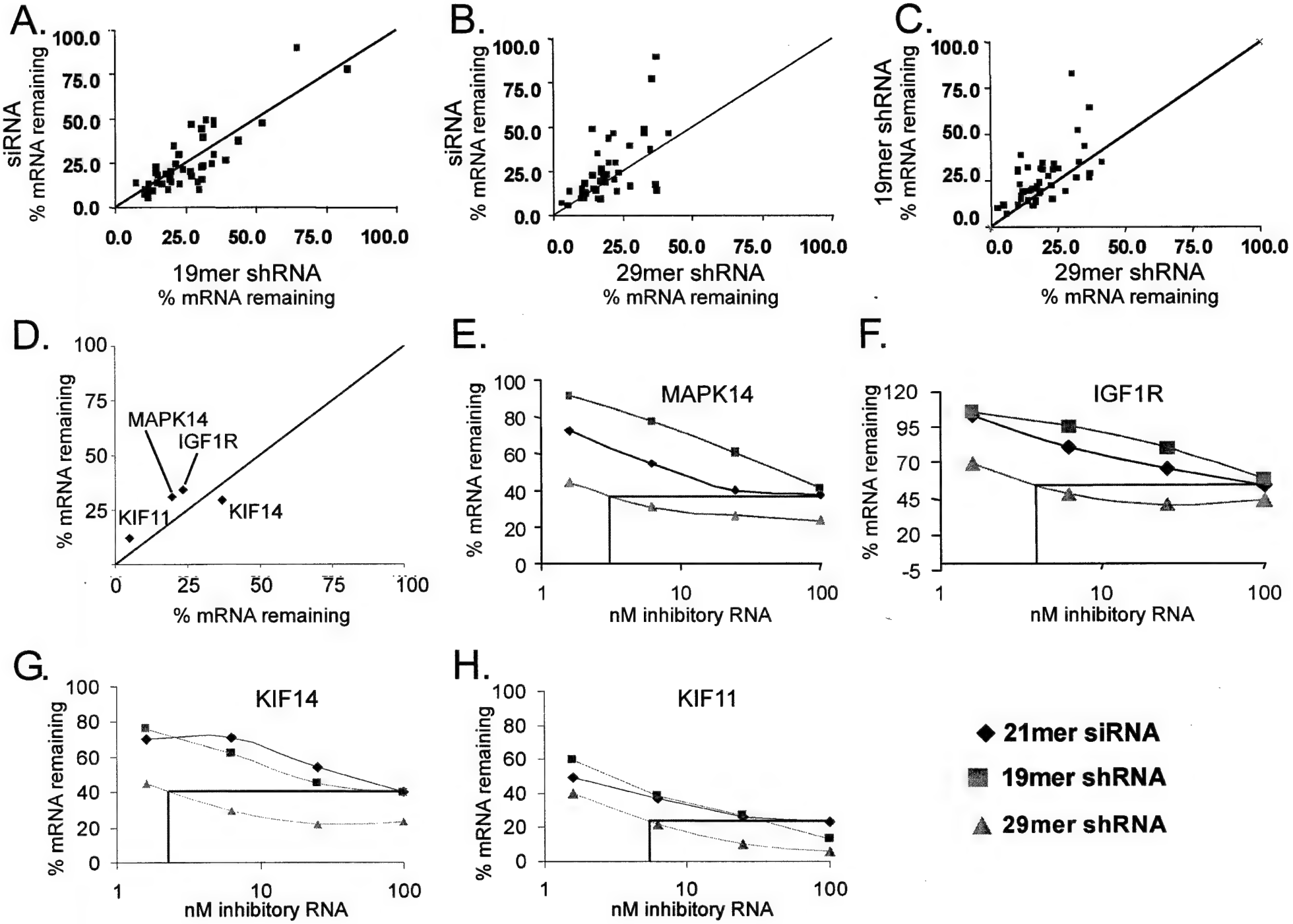
1. Hamilton, A.J. & Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants [see comments]. *Science* **286**, 950-952 (1999).
2. Zamore, P.D., Tuschl, T., Sharp, P.A. & Bartel, D.P. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25-33 (2000).
3. Hammond, S.M., Bernstein, E., Beach, D. & Hannon, G.J. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293-296 (2000).
4. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363-366. (2001).
5. Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R. & Hannon, G.J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146-1150. (2001).
6. Tuschl, T., Zamore, P.D., Lehmann, R., Bartel, D.P. & Sharp, P.A. Targeted mRNA degradation by double-stranded RNA in vitro. *Genes Dev* **13**, 3191-3197 (1999).
7. Song, J.J., Smith, S.K., Hannon, G.J. & Joshua-Tor, L. Crystal Structure of Argonaute and its implications for RISC slicer activity. *Science* **Published Online 7/29/04** (2004).
8. Liu, J. et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **Published online 7/29/04** (2004).
9. Caplen, N.J., Parrish, S., Imani, F., Fire, A. & Morgan, R.A. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci U S A* **98**, 9742-9747. (2001).
10. Elbashir, S.M. et al. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498. (2001).

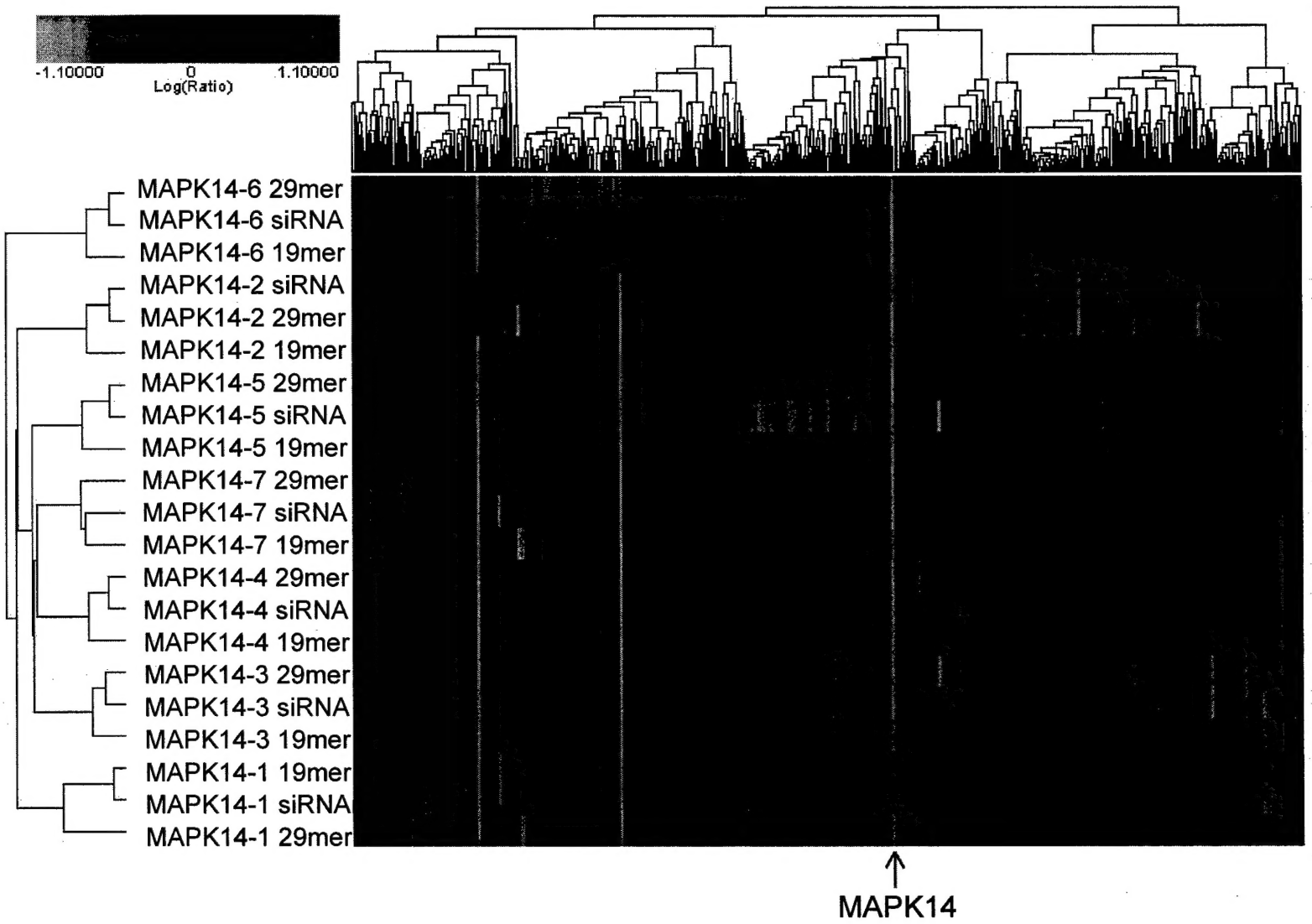
11. Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W. & Tuschl, T. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *Embo J* **20**, 6877-6888. (2001).
12. Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297 (2004).
13. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419 (2003).
14. Hutvagner, G. et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834-838. (2001).
15. Ketting, R.F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev* **15**, 2654-2659. (2001).
16. Grishok, A. et al. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106**, 23-34. (2001).
17. Knight, S.W. & Bass, B.L. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269-2271. (2001).
18. Brummelkamp, T.R., Bernards, R. & Agami, R. A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. *Science* **21**, 21 (2002).
19. Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J. & Conklin, D.S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* **16**, 948-958. (2002).
20. Zeng, Y., Wagner, E.J. & Cullen, B.R. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* **9**, 1327-1333. (2002).
21. Sui, G. et al. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc Natl Acad Sci U S A* **99**, 5515-5520. (2002).
22. Lee, N.S. et al. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* **20**, 500-505. (2002).
23. Paul, C.P., Good, P.D., Winer, I. & Engelke, D.R. Effective expression of small interfering RNA in human cells. *Nat Biotechnol* **20**, 505-508. (2002).
24. Lee, R.C. & Ambros, V. An extensive class of small RNAs in Caenorhabditis elegans. *Science* **294**, 862-864. (2001).
25. Lau, N.C., Lim, L.P., Weinstein, E.G. & Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science* **294**, 858-862. (2001).
26. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858. (2001).
27. Schwarz, D.S. et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208 (2003).

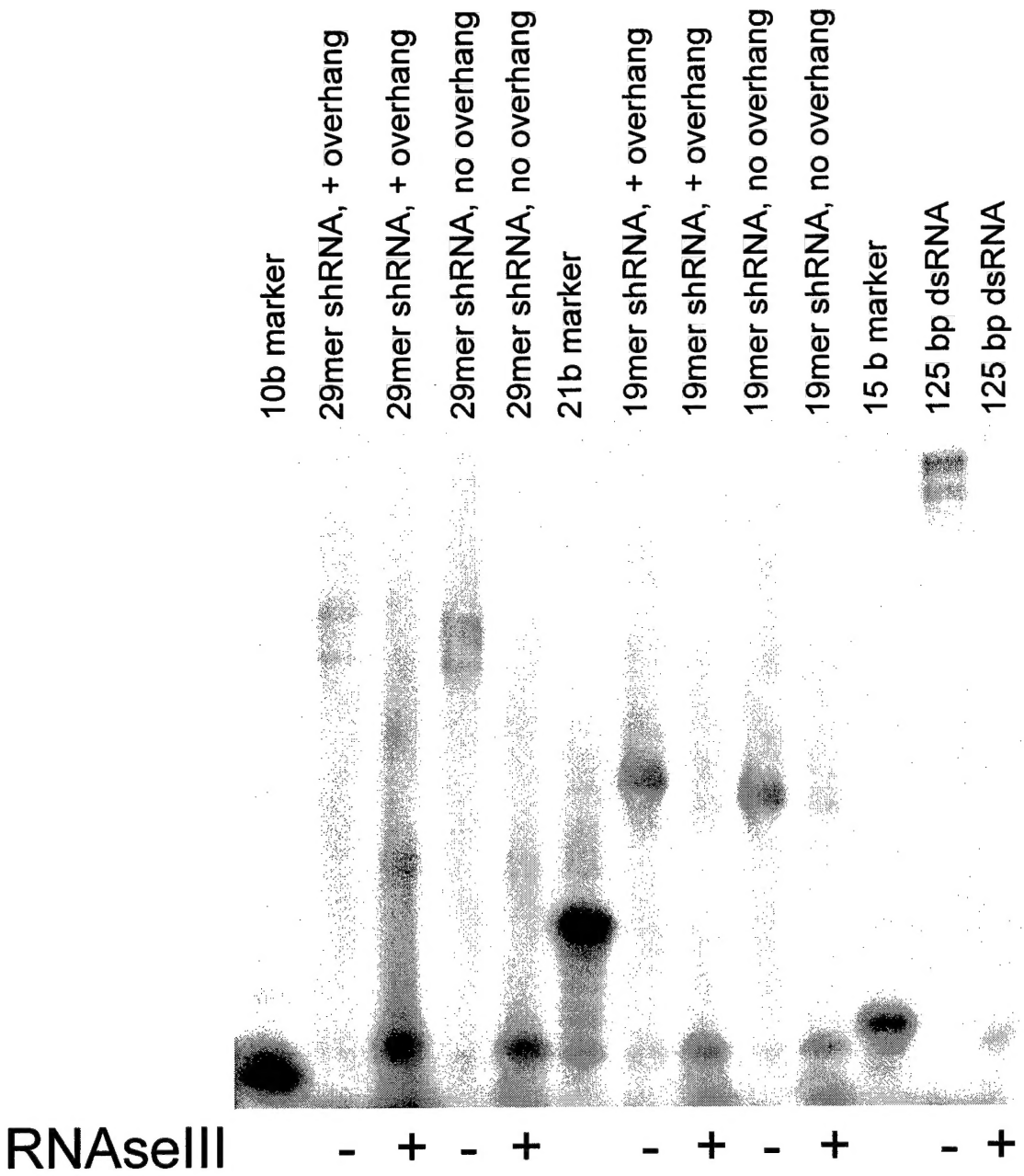
28. Silva, J.M., Sachidanandam, R. & Hannon, G.J. Free energy lights the path toward more effective RNAi. *Nat Genet* **35**, 303-305 (2003).
29. Khvorova, A., Reynolds, A. & Jayasena, S.D. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-216 (2003).
30. Lee, Y.S. et al. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69-81 (2004).
31. Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W. & Sontheimer, E.J. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83-94 (2004).
32. Tomari, Y. et al. RISC assembly defects in the Drosophila RNAi mutant armitage. *Cell* **116**, 831-841 (2004).
33. Zhang, H., Kolb, F.A., Brondani, V., Billy, E. & Filipowicz, W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo J* **21**, 5875-5885. (2002).
34. Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. & Kutay, U. Nuclear export of microRNA precursors. *Science* **303**, 95-98 (2004).
35. Ma, J.B., Ye, K. & Patel, D.J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318-322 (2004).
36. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**, 576-577 (2004).
37. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426**, 465-469 (2003).
38. Song, J.J. et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**, 1026-1032 (2003).
39. Yan, K.S. et al. Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468-474 (2003).
40. Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57-68 (2004).
41. Jackson, A.L. et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**, 635-637 (2003).
42. Doi, N. et al. Short-Interfering-RNA-Mediated Gene Silencing in Mammalian Cells Requires Dicer and eIF2C Translation Initiation Factors. *Curr Biol* **13**, 41-46. (2003).
43. Hughes, T.R. et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* **19**, 342-347 (2001).

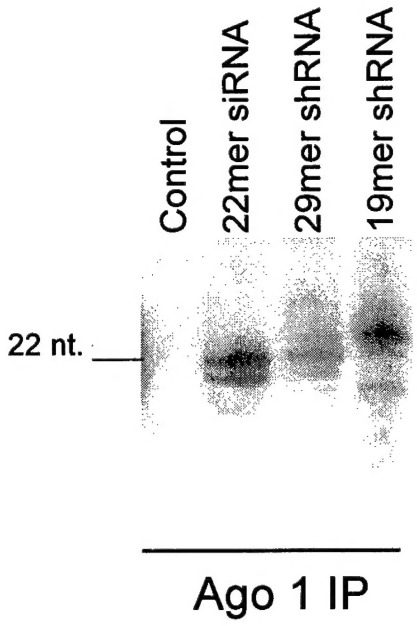

A.
19mer shRNA
NNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNN
CC
AA
19mer shRNA with 2 nucleotide 3' overhang
NNNNNNNNNNNNNNNNNNN
UUNNNNNNNNNNNNNNNNNN
CC
AA
29mer shRNA
NNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNN
CC
AA
29mer shRNA with 2 nucleotide 3' overhang
NNNNNNNNNNNNNNNNNNNNNNNNNNNNN
UUNNNNNNNNNNNNNNNNNNNNNNNNNNNN
CC
AA
Luciferase 29mer
AGUUGCGCCCGCGAAUGAUAUUUAUAAUG
B.
20 nt marker
19mer shRNA
overhang
22 nt marker
29mer shRNA
overhang
21 nt marker
500 bp. dsRNA
500 bp. dsRNA
- - + +
+ + - -
+ - - + + - - + + -
Dicer
C.
29mer shRNA
overhang
22 nt marker
19mer shRNA
overhang
500 bp. dsRNA
500 bp. dsRNA
- - ++
+ + - -
- + - + - + - + + -
Dicer


A.
Primer only
21 nt marker
29mer shRNA in vitro
22 nt marker
29mer shRNA in vivo
Ago 2 IP
Ago 1 IP
20 nt marker
22 nt siRNA
22 nt siRNA
21 nt siRNA
21 nt siRNA
25 nt marker
RT + - + - + - + + - + - +
B.
21mer siRNA
21mer siRNA
22mer siRNA
22mer siRNA
20 nt. marker
29mer shRNA
29mer shRNA
22nt. marker
Ago1 IP
Ago1 IP
Ago2 IP
Ago2 IP
Primer only
15nt. marker
25nt. marker
RT - + - + + - + - + - +

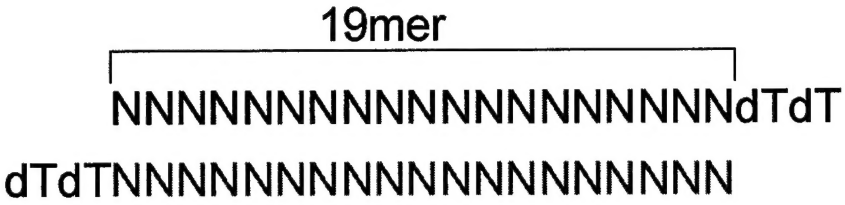

A.
siRNA
% mRNA remaining
19mer shRNA
% mRNA remaining
B.
siRNA
% mRNA remaining
29mer shRNA
% mRNA remaining
C.
19mer shRNA
% mRNA remaining
29mer shRNA
% mRNA remaining
D.
% mRNA remaining
MAPK14
IGF1R
KIF11
KIF14
% mRNA remaining
E.
MAPK14
% mRNA remaining
nM inhibitory RNA
F.
IGF1R
% mRNA remaining
nM inhibitory RNA
G.
KIF14
% mRNA remaining
nM inhibitory RNA
H.
KIF11
% mRNA remaining
nM inhibitory RNA
21mer siRNA
19mer shRNA
29mer shRNA


-1.10000
0
1.10000
Log(Ratio)
MAPK14-6 29mer
MAPK14-6 siRNA
MAPK14-6 19mer
MAPK14-2 siRNA
MAPK14-2 29mer
MAPK14-2 19mer
MAPK14-5 29mer
MAPK14-5 siRNA
MAPK14-5 19mer
MAPK14-7 29mer
MAPK14-7 siRNA
MAPK14-7 19mer
MAPK14-4 29mer
MAPK14-4 siRNA
MAPK14-4 19mer
MAPK14-3 29mer
MAPK14-3 siRNA
MAPK14-3 19mer
MAPK14-1 19mer
MAPK14-1 siRNA
MAPK14-1 29mer
MAPK14

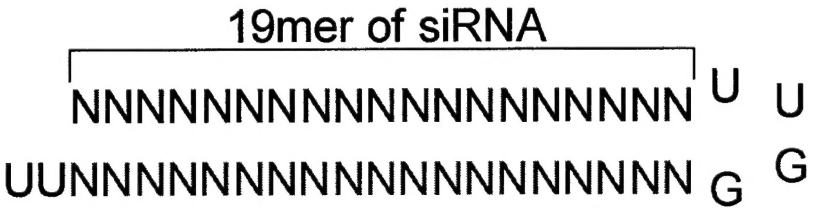

10b marker
29mer shRNA, + overhang
29mer shRNA, + overhang
29mer shRNA, no overhang
29mer shRNA, no overhang
21b marker
19mer shRNA, + overhang
19mer shRNA, + overhang
19mer shRNA, no overhang
19mer shRNA, no overhang
15 b marker
125 bp dsRNA
125 bp dsRNA
RNAseIII
- + - + - + - + - + - +


Control
22mer siRNA
29mer shRNA
19mer shRNA
22 nt.
Ago 1 IP

siRNAs



Synthetic 19mer shRNAs



Synthetic 29mer shRNAs